

MODELING OCCURRENCE OF THE GREEN SALAMANDER, ANEIDES  
AENEUS, IN THE BLUE RIDGE ESCARPMENT

A thesis presented to the faculty of the Graduate School of  
Western Carolina University in partial fulfillment of the  
requirements for the degree of Master of Science in Biology.

By

Rebecca Hale Hardman

Director: Dr. Joseph Pechmann  
Associate Professor of Biology  
Biology Department

Committee Members: Dr. Thomas Martin, Biology  
Dr. Ronald Davis, Natural Resources

October 2014

## ACKNOWLEDGEMENTS

I would like to thank everyone who helped me develop my question and determine how I was ever going to answer it. Thanks to my committee for all their support and special thanks to Bill Sutton.

## TABLE OF CONTENTS

List of Tables .....	iv
List of Figures .....	v
List of Abbreviations .....	vi
Abstract .....	vii
Introduction.....	9
Methods.....	14
Study Area .....	14
Model Building .....	16
Species Location Datasets.....	16
Environmental Datasets .....	17
Model Evaluations .....	18
Field Surveys .....	19
Results.....	21
Model Evaluations .....	21
Field Surveys .....	26
Discussion.....	29
Literature Cited .....	34

## LIST OF TABLES

Table	Page
1. Model Evaluations .....	23
2. Variable Contributions.....	25

## LIST OF FIGURES

Figure	Page
1. Study Area .....	15
2. Models with Binary Outputs .....	22
3. Models with Binary Outputs and Overlaying <i>Aneides aeneus</i> Presence Points. ....	24
4. Survey Locations .....	26

## LIST OF ABBREVIATIONS

MaxEnt.....	Maximum Entropy
LR .....	Logistic Regression
GRG .....	Green River Game Lands
DSP .....	Dupont State Park
SMS .....	South Mountains State Park
TCP .....	Transylvania County Private Lands

## ABSTRACT

### MODELING OCCURRENCE OF THE GREEN SALAMANDER, *ANEIDES AENEUS*, IN THE BLUE RIDGE ESCARPMENT

Rebecca Hardman

Western Carolina University (March 2011)

Director: Dr. Joseph Pechmann

Amphibian species have experienced global declines since the 1970s and plethodontid salamanders are no exception. The green salamander, *Aneides aeneus*, is a plethodontid salamander that has experienced declines throughout its range in the Blue Ridge Escarpment.

Species distribution models are algorithms that predict occurrences of a species across a landscape and can be used to determine conservation priority areas. However, there are commonly only presence locations without corresponding absence locations available to a researcher. These presence-only datasets can present a challenge when trying to depict reliable distributions for a species of concern. Maximum Entropy (MaxEnt) is an algorithm empirically tested to model species distributions given presence-only datasets.

I used landscape-level species distribution models including MaxEnt and logistic regression to model the occurrence of green salamanders across the Blue Ridge Escarpment of North Carolina. These models were used to assess particular features associated with *A. aeneus* presence as well used to search for new localities.

MaxEnt models outperformed logistic regressions for all methods of evaluation. MaxEnt models had fairly low omission (false negative) and commission (false positive) rates whereas my logistic regression had extremely high error rates for both. “Area Under the Receiver Operator Curve” evaluation scores were excellent (0.96) and good (0.81) for the top Maxent model and logistic regression, respectively.

*Aneides aeneus* is known to be associated with habitat that includes rock outcroppings with thin, deep crevices. My models indicated that forested areas, intermediate elevations, and shallow soils of particular types are desirable landscape features for *A. aeneus*. Soil was the most important variable in all models, accounting for almost half of the variation in MaxEnt models. Elevation accounted for most of the remaining variation. Percent canopy cover accounted for 4-6.5% of the variation in Maxent models. While these models did not specifically predict presence of outcrops, they were extremely helpful in identifying habitat with conditions supportive for *A. aeneus* if a rock outcrop was present. With the help of these models I discovered one previously unknown locality for *A. aeneus* and am confident additional locations can be found.



## INTRODUCTION

Amphibians have experienced strong global declines since the 1970's; an estimated 33% of all extant amphibian species are currently in decline (Stuart et al. 2004; Wake and Vrendenberg 2008). Some plethodontid salamanders have similarly experienced large scale declines (Highton 2005; Rovito et al. 2009), and while this group is very diverse and abundant in the Appalachian region, many species are vulnerable to threats such as urbanization, climate change, and air and water pollution (Petranka et al. 1993; Petranka 1998). However, the elusive nature of many salamander species makes it difficult to obtain accurate population estimates and ultimately to determine their longterm conservation status.

The green salamander (*Aneides aeneus*) is a plethodontid reported to have had noticeable population declines throughout the Blue Ridge Escarpment since the mid-1970's (Snyder 1983; Corser 2001). It is listed as a species of concern in every state throughout its current range, and is listed as endangered by the state of North Carolina. Habitat requirements therefore need to be well understood to implement proper long term conservation and management plans. *Aneides aeneus* is commonly found in moist, but not wet, rock crevices in shaded mesic forest habitat where the temperatures are lower than that of the surrounding area throughout the summer (Netting and Richmond 1932; Gordon 1952; Bruce 1968). Anatomical adaptations, such as dorsoventrally flattened bodies, facilitate easy maneuvering within these crevices. The presence of thin, deep crevices in the rock faces along with other attributes, rather than rock-type itself, has been described as one of the most accurate indicators of *A. aeneus* populations (Gordon 1952;

Reed et al. 2009). Relative abundances of some populations is also correlated with the presence of certain hardwood tree species, especially American beech (*Fagus grandifolia*). This is likely due to *Aneides* use of the abundant moist crevices on the trunks of these trees (Wilson 2003; Waldron & Humphries 2005). Dispersal between outcrops is suspected, but this has been difficult to corroborate. Populations reach their highest altitudes within the Blue Ridge Escarpment, between 500 and 1300 m (Corser 2001), whereas populations in the Cumberland Plateau and Allegheny Mountain geographic provinces are found up to 915 m (Pauley et al. 1993).

While the above mentioned studies have determined some microhabitat variables important to *A. aeneus*, few have examined habitat associations on a landscape level. Landscape scale environmental variables can be modeled and mapped using Geographic Information Systems (GIS) to further determine the ecological requirements of *Aneides aeneus*. This would potentially permit the identification of critical habitats throughout the range of *A. aeneus* and identification of previously unknown populations more quickly than by strictly field-based surveys (Guisan et al. 2013).

Hutchinson (1957) described a niche as an n-dimensional hyper-volume where a species's requirements fall within a unique set of parameters of biotic and abiotic factors corresponding to each dimension. Modeling the conditions under which organisms are found can help demonstrate the realized niche of the species.

Species Distribution Models are the statistical evaluation of the distribution of a species, its relationships with environmental variables, and the use of these to predict occurrences in other geographic and temporal spaces (Franklin 2009). Although true fundamental niches cannot be clearly determined based on individual presences along

environmental gradients alone, these models are still extremely useful for examining environmental associations which may ultimately affect species distributions and prediction of high quality habitat.

Choosing among the many available approaches for species distribution modeling depends on the researcher's goals and available resources (Aguirre-Gutiérrez et al. 2013). Each approach has different assumptions, and it is important to choose a model that represents the ecological reality of the organism (Austin and Gaywood 1994; Guisan and Thuiller 2005). Guisan and Thuiller (2005) also stress the importance of applying ecological theory for choosing the right predictor variables. Fishing through unnecessary variables can decrease model power, resulting in an over-fit model. Therefore, it is imperative for a researcher to have an *a priori* understanding of which ecological factors are most likely to affect a particular organism. Beyond developing biologically relevant hypotheses, it is also imperative to select the right spatial scale for data layers when dealing with a spatial GIS model (Pandit et al. 2010). Larger pixel sizes lose important habitat details by combining adjacent values, and decrease the number of available sample points.

Only one prior study has modeled *Aneides aeneus* habitat suitability (Lipps 2005). This study created a predictive GIS model for *A. aeneus* in two Ohio counties and helped increase knowledge of new sites and *A. aeneus* habitat. However, no studies have modeled habitat for *A. aeneus* over an entire region using statistical models commonly used for species distribution modeling. Identification of suitable habitat over a large scale can be useful for identifying distributional limits, developing conservation strategies, and identifying critical habitat for acquisition.

Traditionally both presence and absence datasets were thought to be necessary to correctly model habitat suitability. This requirement poses a problem for many possible studies as absence data requires a lot more research, may be unreliable depending upon the detection probability of the organism, or may be unavailable. Regressions require absence data alongside presence data and therefore pseudo-absences often must be produced. These pseudo-absences are randomly created either based loosely off surveys of other species or out of the general background itself. They are not true absences because they are not survey verified but at least may represent random background available habitat in contrast to known presences. Results are affected since some may not be true absences and their creation was not done in the same manner as that of the presences (Zaniewski et al. 2002; Engler et al. 2004; Elith and Leathwick 2007).

Logistic regression is a generalized linear model commonly used to create species distribution models of a fitted binomial distribution. It is an easy algorithm to use and has been effective in modeling certain distributions. Though still in use, this model has fallen out of favor when compared to newer techniques, especially due to absence dataset requirements. New methods are often compared against it.

Over the past decade newer modeling techniques have emerged that have proved to be as effective as older techniques but are able to use presence-only data. Maximum Entropy (MaxEnt) is one in wide use (Phillips 2006). MaxEnt has historically tested well against both presence-only and presence/absence modeling techniques when predicting species distributions (Peterson 2007; Elith et al. 2006; Rebelo and Jones 2010). While there is evidence of pitfalls to MaxEnt, most are revealed to be problems inherent in

presence datasets and misinterpretation of outputs, rather than in the model algorithm itself (Yackulic et al. 2013).

MaxEnt has been recently shown to be mathematically equivalent to a Poisson regression, identifying shared assumptions to other modeling techniques like logistic regression (Renner and Warton 2013). Nevertheless, MaxEnt continues to perform and create species distribution models that stand up well in comparison with logistic regressions and newer algorithms (Aguirre-Gutiérrez et al. 2013).

This study evaluated landscape-level habitat requirements of the green salamander within the Blue Ridge Escarpment and created a map of predicted occurrences across this region by using MaxEnt species distribution modeling and a more empirically tested GLM. My goal was to facilitate the location of more sites occupied by *Aneides aeneus* and to obtain a better understanding of its habitat requirements.

## METHODS

### *Study Area:*

My study area included all or part of ten counties in the southern Blue Ridge Mountains of western North Carolina. These were (from west to east): Macon, Jackson, Haywood, Transylvania, Buncombe, Henderson, Polk, Rutherford, McDowell, and Burke counties (Figure 1). The western part of the study area included the Nantahala and Cowee Mountains in Macon and Jackson counties, moving northeast to the Great Balsams, Newfound, and Great Craggy Mountains in Haywood and Buncombe counties. The eastern end of the region drops precipitously along the edge of the Blue Ridge escarpment. The Eastern Continental Divide slices from north to south along the eastern edge of the region and then along the South Carolina border, keeping most watersheds turning west into the Tennessee River drainage.

The Nantahala and Pisgah National Forests encompass much of the forested area. Dupont State Forest and Gorges State Park encompass other forests in Transylvania County. South Mountains State Park reaches into the eastern part of the study area. Intense historical land use is a dominant theme across the landscape with much of the current forest having re-grown from clear-cutting for timber harvest in the early 1900's. Mid-elevation upland habitats consist of mixed hemlock and hardwood forests but are changing due to the hemlock woolly adelgid (Eschtruth et al. 2006). Seeps and first and second order streams bordered by *Rhododendron sp.* are common under the canopy.

Rock outcroppings in the Blue Ridge escarpment are not as large and continuous as those found in the Cumberland Plateau and typically consist of small to large boulders

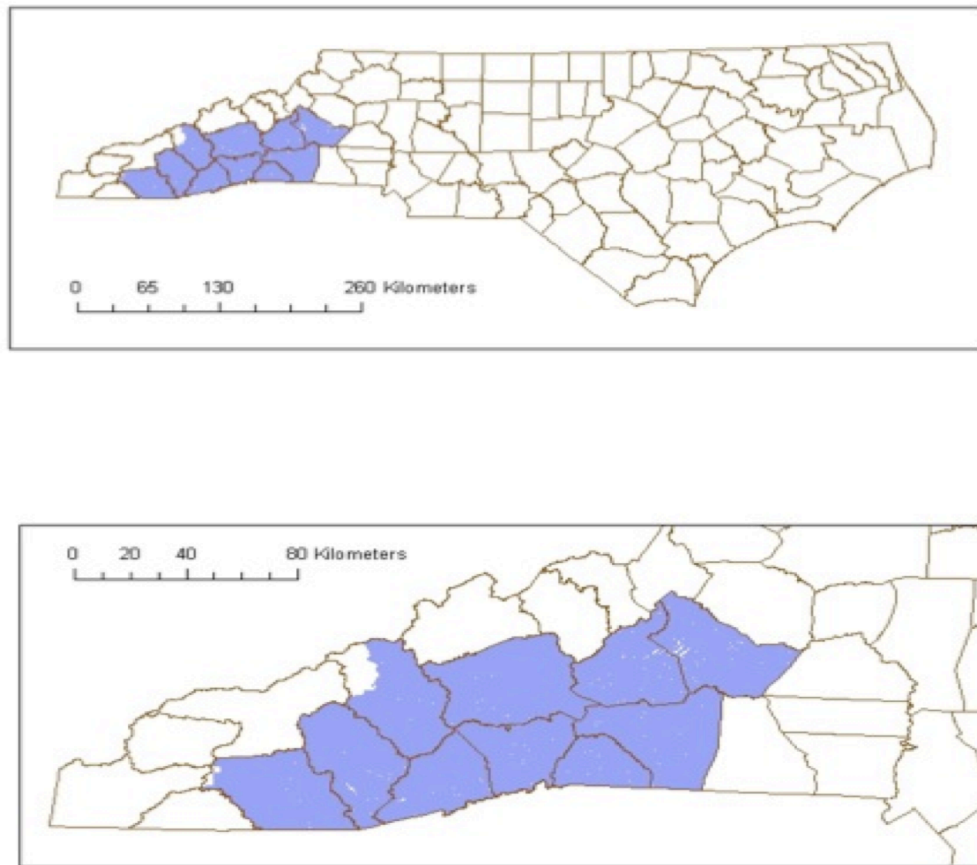


Figure 1. Study Area. Shaded in light blue is area denoting 10 counties in western North Carolina used for study.

scattered randomly in the landscape. The exact locations of outcroppings are hard to predict because they are hidden underneath the canopy and small in both length and in height.

### *Model Building:*

I used two modeling techniques to estimate species distribution for *Aneides aeneus*: a logistic regression (GLM with binomial distribution) and a Maximum Entropy (MaxEnt) model. I generated a logistic regression using SPSS version 9.0 (SPSS, Inc., Chicago IL) from values extracted at presence and pseudo-absence points. For MaxEnt models, my entire study area background was analyzed against the values at each presence point using MaxEnt Species Distribution Modeling Software (Phillips 2006).

All models were converted to predictive maps across western North Carolina in the same set of counties from which the models were created to avoid any changes in correlations of the predictor variables that are likely to occur when models project onto a spatially and/or temporally distant region (Austin 2002). For MaxEnt I used the ASCII output created by the program, and for the GLM I applied the logistic output equation in the ArcMap function “raster calculator” (ArcGIS v. 10.1 ESRI, Inc., Redlands, CA)

### *Species Location Datasets*

Logistic regression and MaxEnt models were based on presence data for North Carolina obtained from archives of the NC Wildlife Resources Commission and NC Natural Heritage databases. I converted any location data in the form of a polygon to a point, converting its geometric center to a centroid coordinate.

I accounted for spatial autocorrelation, which has been shown to be a major component of MaxEnt model quality (Syfert et al. 2013). Dispersal is considered to be limited in the Blue Ridge escarpment with only occasional findings far from outcrops or cliff (Riedel et al. 2006). I created a threshold of 30 m as a minimum distance between presence points. This value was based on the size of several raster layers available with a



minimum pixel size of 30 m, and the distance considered large enough to encompass daily movements (Miloski 2010) and keep sites independent. This scale is much smaller than many studies at 1 km resolution and may facilitate detection of some variation in microclimate important to a lungless salamander. I randomly chose only one out of any group of points within the 30m threshold distance to be used in my models.

For the logistic regression, I generated pseudo-absence points which were randomly spaced within the boundaries of my study area. Pseudo-absences should be well distributed within the region being modeled (Lobo et al. 2010). I weighted pseudo-absences the same as presences which has been shown to produce more accurate models (Barbet-Massin et al. 2012). For MaxEnt I chose the default for the entire background of the study area to be used against the presence locations.

#### *Environmental datasets*

I acquired a variety of landscape variables in GIS raster format of pixel size 30 m or less. I included percent canopy cover (PCC) from Landsat (<http://landsat.gsfc.nasa.gov>). I included canopy height and elevation from North Carolina Digital Elevation Models( NC DEM; [www.ncdot.gov](http://www.ncdot.gov)), both previously created from LiDAR point data. From elevation, I generated both slope and aspect. Lastly, I incorporated NRCS soils data (<http://soildatamart.nrcs.usda.gov/>), either as a raw NRCS soil-type category or converted to the attribute of soil depth-to-bedrock using Soil Data Viewer (<http://www.nrcs.usda.gov/>).

For logistic regression and MaxEnt analyses both aspect and soil were converted into continuous variables, the former based on eastness and northness, the latter based on depth-to-bedrock. Eastness and northness represent angular distance from respective

compass directions, where for eastness, east ( $180^\circ$ ) is 1, West ( $270^\circ$ ) is 0, and all directions are values in between. For the logistic regression, I transformed elevation from true elevation to distance to most frequent elevation of 800m for *Aneides aeneus* presence. All raster layers were standardized to contain identical pixels at 30 m resolution. MaxEnt can include categorical variables with several categories in the analysis and therefore soil-type category was used for two models without being transformed.

MaxEnt and species distribution models are negatively affected by either excessively complex or overly-simple models (Warren and Seifert 2011). I considered six environmental covariates to be enough to avoid either extreme. Pearson correlations determined no collinearity ( $r > 0.80$ ) among the variables used.

#### *Model Evaluations:*

To compare models I calculated both AUC [Area Under the Receiver Operator Curve (ROC) Curve] using randomly selected 30% hold-out test locations as well as omission (false negative) and commission (false positive) rates (Raes and ter Steege 2007).

Omission rates determine the sensitivity of the model (higher omission results in decreased sensitivity) and commission rates determine specificity (higher commission result in lower specificity). AUC is a common method for evaluation of continuous output models from presence-only data (Raes and ter Steege 2007). An ROC is created by an iterative process of incremental changes to the threshold value of a model. Sensitivity and “1 – specificity” are plotted against each other for each threshold value. A value between 0-1 (0-100% of area possible) is obtained where the greater the area under

the curve, the better predictive ability the model should have. Models with an AUC >0.5 predict better than random.

AUC values have historically been shown to be inflated in certain situations, e. g., when species distributions are over small areas relative to the background. They can give false confidence in sometimes subpar models and incorporating more than one method for model evaluation is recommended (Warren and Seifert 2011). I, therefore, calculated omission rate as percent of true presences predicted as absent, and commission rates as the percent of pseudo-absence background points predicted as present. Evaluating these required a binary output of either predicted present or absent. I used a threshold value of maximized sum of sensitivity and specificity which is a well supported method to create binary results from continuous outputs (Liu et al. 2005; Jimenez-Valverde and Lobo 2006).

For the logistic regression I considered a relationship significant for any variables with a p-value < 0.05. For MaxEnt I determined variable importance using a jackknifing approach to changes in AUC and from program calculations of variable % contribution. I then compared changes in variable importance between the modeling approaches.

In preliminary models I noted a strong contribution of the soils variable and subsequently created an additional model with soil category as the only predictor variable.

#### *Field Surveys:*

Ground-truthing surveys occurred during spring through fall of 2010 and fall of 2011. Survey locations were selected based on two criteria: 1) high (> 0.85) predicted presence on output maps, and 2) areas with no previously known presence locations.

Given these characteristics we focused on four main areas of western North Carolina: Transylvania county private property, Green River Games Reserve, Dupont State Park, and South Mountains State Park. Within these regions I searched a total of thirty sites. Sites were defined as an area encompassing a 90 meter radius around one previously determined coordinate. I conducted all surveys during the day and focused efforts specifically on rock outcrops. I also searched the bark of beech trees which have been illustrated to be important to *A. aeneus* (Waldron and Humphries 2005). Surveys can be easily biased depending on the activity of the organism outside the crevice which is heavily influenced by season and time of day (Waldron and Humphries 2005) as well as proximity to rainfall events (Humphries pers comm). I attempted to limit surveys to early morning, cooler, or cloud covered days which are conditions that maximize detection on and around crevices.

## RESULTS

### *Model Evaluations*

I had four different models for comparison: (1: Logistic Regression (LR), 2: MaxEnt model with identical covariates and presences to LR, 3: MaxEnt model with same covariates as LR except soil as soil type instead of depth-to-bedrock, and 4: MaxEnt using soil type as the only predictor. Models created using MaxEnt had extremely good predictive ability evaluated with both AUC and omission/ commission rates (Table 1). AUC scores for MaxEnt models ranged between 0.925 and 0.967. The soil-only MaxEnt model had the best AUC score (AUC= 0.967). The logistic regression (AUC= 0.815, Table 1) fell behind all MaxEnt models but within the range to be considered a reasonably good discriminatory model of between 0.7 and 0.9 (Pearce and Ferrier 2000).

Application of omission and commission rates to evaluation separated MaxEnt and logistic regression models even further. After converting models to a binary output based on maximum sensitivity and specificity, MaxEnt models exhibited the lowest omission and commission rates. Again, the logistic regression fared the worst with commission/omission error rates of 0.93 and 0.41 respectively. Low error rates of all MaxEnt models equated to all having higher sensitivity and specificity than the logistic regression (Table 1). Model 3 (MaxEnt with all variables and soil-type) had the lowest error rates. Pseudo-absences used to calculate commission errors were originally incorporated into the logistic regression as true absences, yet that model had much higher

commission than any of the MaxEnt models. MaxEnt models did not use those points in model building.

Area predicted as present within the study was dramatically different among most models. The logistic regression predicted a large scattered range that does not follow the known range of *A. aeneus*. It had the largest area of predicted presence in the study

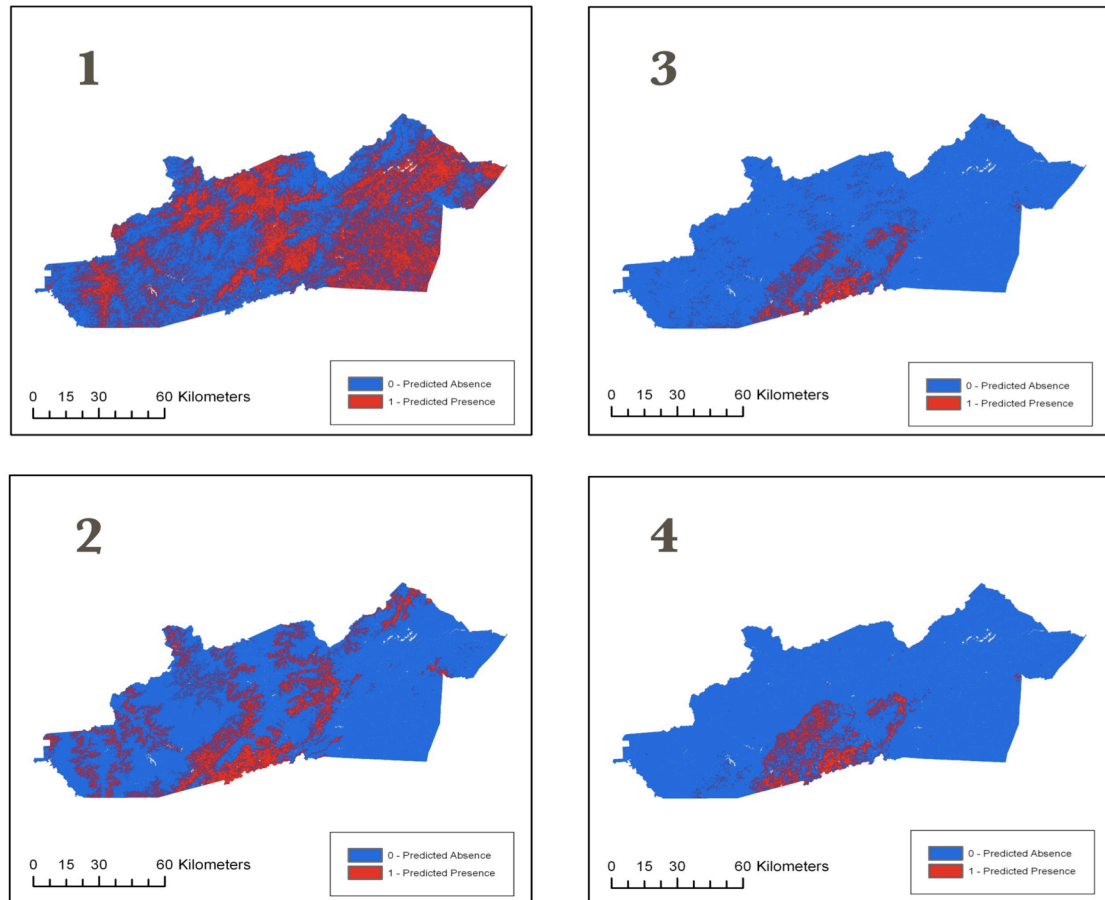


Figure 2. Models with binary outputs. Thresholds determined based on maximum sensitivity + specificity. Models are as follows: 1. Logistic Regression 2. MaxEnt with all continuous variables with soil as depth-to-bedrock 3. MaxEnt with all variables and soil as soil-type 4. MaxEnt with soil-type as only covariate.

Table 1. Model Evaluations. Evaluations for all four models are listed below. AUC represent area under the curve values. Area predicted represents total area for positive predicted presence within the study limits after each model was converted to a binary output. Error rates and sensitivity(Sens)/specificity(Spec) were all calculated on the binary outputs. Omission represents amount of true presence predicted as false and commission amount of pseudo-absences predicted as present. Sensitivity and specificity directly relate to former values where sensitivity and specificity increase with respective decreases of omission and commission.

Model No.	Algorithm	Covariates	AUC	Area predicted (km2)	Omission	Commission	Sens	Spec
1	LR	All continuous variables; Soil as depth to bedrock	0.815	1957	0.93	0.41	0.06	0.59
2	MaxEnt	All continuous variables; Soil as depth to bedrock	0.925	963	0.08	0.16	0.92	0.84
3	MaxEnt	All continuous variables except soil as soil type	0.958	728	0.08	0.06	0.92	0.94
4	MaxEnt	Soil type only	0.967	725	0.12	0.06	0.88	0.94

region (Table 1, Figs 2 and 3). MaxEnt models 3 and 4 that use soil as a categorical variable predict the smallest area but in contrast have more overlap with the known *A. aeneus* range (Table 1, Figs. 2 and 3). Maxent model 2 predicts an area between that of the logistic regression and MaxEnt models 3 and 4, with a predicted range that extends much further north than the latter two.

Model 4 was evaluated to have good predictions based on AUC alone, but adding other environmental variables (Model 3) created a more reliable model when evaluated with both AUC and error rates. Evaluations are not as clear when using AUC alone, preventing true comparisons in model interpretation, especially when AUC can often be spuriously inflated (Smith 2013).

Soil was the most important predictor variable in all models. Depth-to-bedrock had a significant negative correlation to presence in the logistic regression ( $p < 0.001$ ), and had a 43% contribution in MaxEnt model 2 (Table 2). Soil-type had an even larger contribution of 57% in model 3. Elevation was the second most important variable in

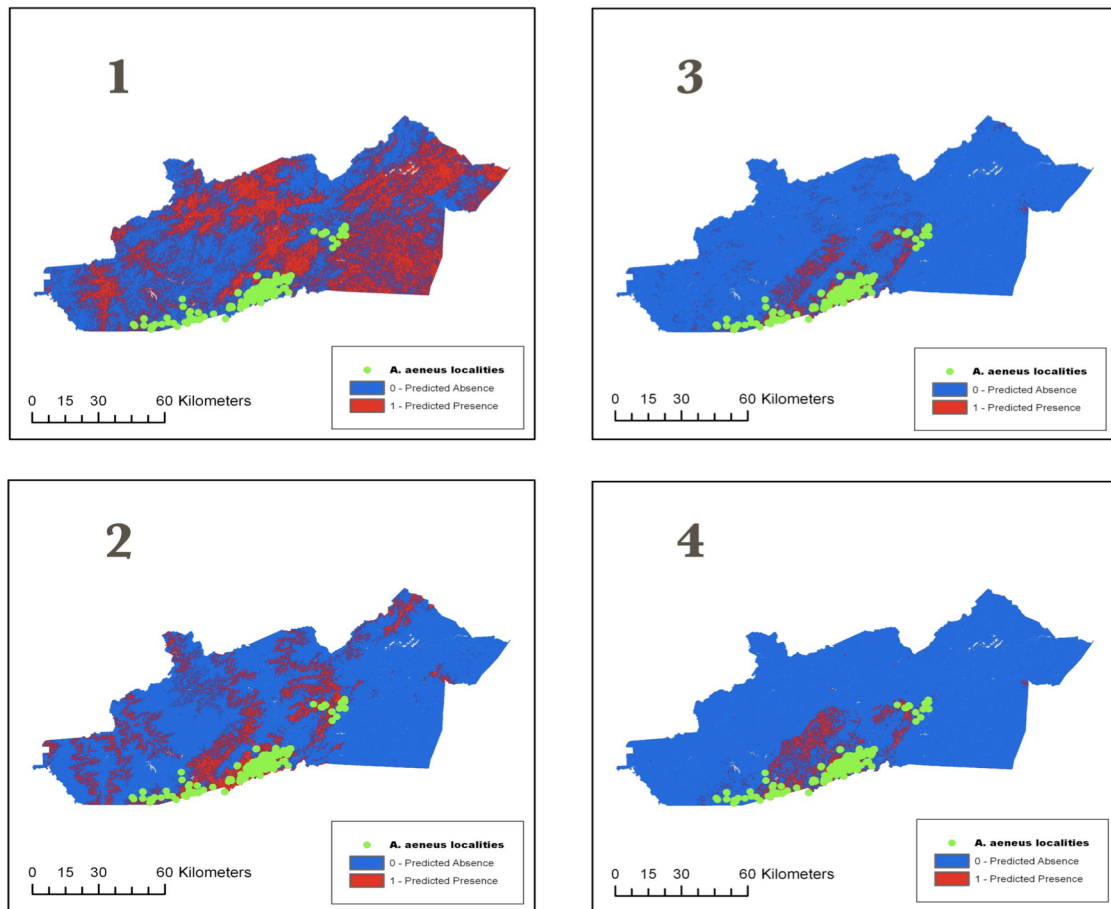


Figure 3. Models with binary outputs and overlaying *Aneides aeneus* presence points. Thresholds determined based on maximum sensitivity + specificity. Models are as follows: 1. Logistic Regression 2. MaxEnt with all continuous variables with soil as depth-to-bedrock 3. MaxEnt with all variables and soil as soil-type 4. MaxEnt with soil-type as only covariate.



MaxEnt models 2 and 3 with 41% and 34% contributions, respectively. Presence was positively correlated with percent canopy cover (PCC), the only other variable with a significant relationship in the logistic regression. Percent canopy cover and all other variables in MaxEnt models had minor (< 7%) contributions. Slope and aspect had the least amount of contribution to any model. These models predict important *A. aeneus*

Table 2. Variable contributions. For the logistic regression (LR) variables with p-values < .05 are denoted as SIG. They are then given a + or – to represent either a positive or negative relationship with presence, respectively. For MaxEnt models, percent contributions are given for each variable. Any variable not used in a given model is denoted with NA (not applied). PCC is % canopy cover.

Variable	LR	MaxEnt 2	MaxEnt 3
Soil Category	NA	NA	57
DTB	<b>SIG -</b>	43	NA
Elevation	non sig	41	34
PCC	<b>SIG +</b>	6.5	4
Canopy Height	non sig	5.5	3
Slope	non sig	2.5	1
Aspect	non sig	1.5	1

habitat to include features of mid elevation, intact canopy and shallow soil types.

Jackknife evaluations of MaxEnt model 2 and 3 showed depth-to-bedrock and soil type most important for high AUCs. However, AUC was highest with all variables included.

Computer evaluations indicate extremely good predictive ability of all MaxEnt models with increasing accuracy when soils were kept in categories. Spatially, the best

models (3 and 4) predicted *A. aeneus* presence along the edge of the Blue Ridge escarpment (Figs 2 & 3).

### *Field Surveys*

While conducting field surveys of high (logistic output  $> 0.85$ ) predicted-probability sites, I noticed a consistent pattern in the landscape of horseshoe-shaped coves or undulating areas with high topographic relief. These sites usually fell within stands of mature forest, however, they did not always contain exposed rock outcroppings. Outcrops were discovered in all four areas targeted for intensive surveys: Green River Game Lands, Dupont State Forest, South Mountains State Park, and on private property in Transylvania County, NC (Fig. 4). No new locations for *A. aeneus* were sighted in the

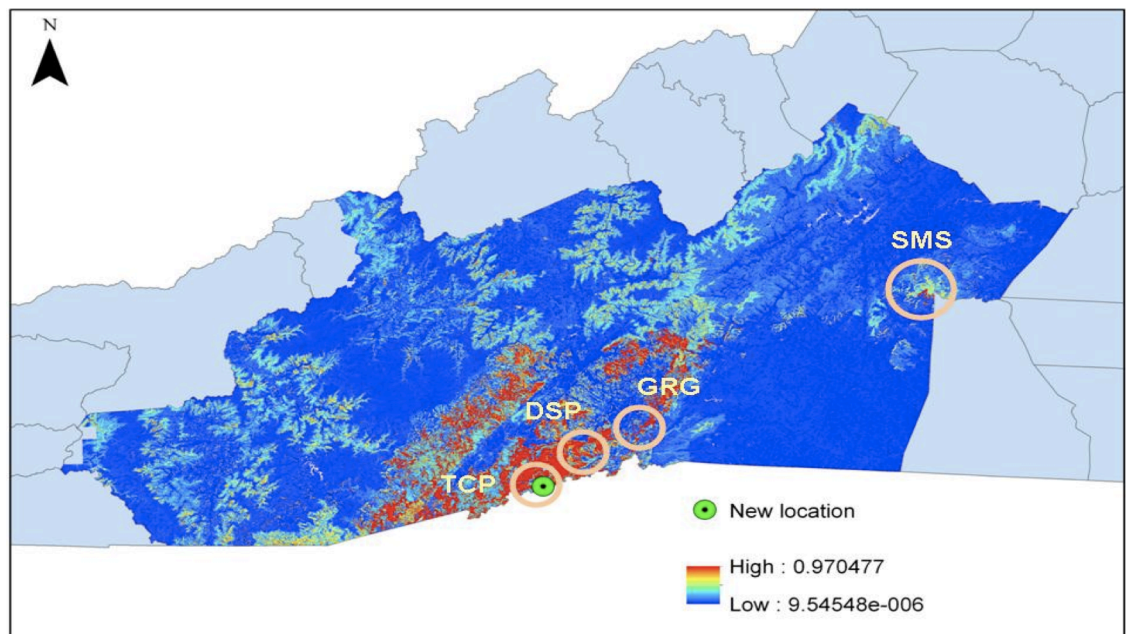


Figure 4. Survey Locations. Survey areas circled over continuous output of Model 3. TCP: Transylvania County private land, DSP: Dupont State Park, GRG: Green River Game Lands, SMS: South Mountains State Park. Newly discovered site denoted by green circle.

first three regions, but I discovered one new site in a large shaded outcrop in Transylvania County.

The Green River Game Lands seemed a promising area to find new sites. This state owned forest was of particular interest because it had several areas with high probability of occurrence, but it also fell inside a large gap between two well documented areas of green salamander presence. I noticed many rock outcrops in areas that seemed to be suitable given landscape attributes, but the outcrops were extremely dusty and dry.

South Mountains State Park was another promising region having areas of predicted presence but no previous records of green salamanders. While it is disjunct from the Blue Ridge escarpment it has very similar topography and interestingly contains a species of plethodontid salamander from the *Plethodon jordani* complex, a taxonomic group centered in the southern Appalachian/ Blue Ridge region. If there has been dispersal of this salamander complex, there could have been dispersal of green salamanders as well. I surveyed along the eastern side of the park and did not locate any animals. Rock outcrops in this area were usually too wet for *A. aeneus*, and I found several *Desmognathus fuscus* which are more common in wetter, seepage-type habitats (Brimley 1944; Petranka 1998). The west end of the park is where my models had larger areas of predicted presence, but because of accessibility issues, I was unable to survey that section.

Dupont State Forest also had high predicted presence according to our top models. Although I did not find new locations there, I only looked in areas where *A. aeneus* had not previously been identified as “present”. This forest is state owned and

falls directly within the known range of the species. It has already been heavily surveyed and monitored and had multiple previously documented populations.

I was able to locate a new population in Transylvania County, and not so coincidentally, it was within my only survey area on private land. This area was in another “hot spot” along the border of North and South Carolina (Fig 4) and contained several scattered tracts of private property for homes, gated communities and camps. Most of the area was similar to previous visited sites in topography and forest stage. Some rhododendron patches were present. Understory composition varied but was predominantly open. I discovered one very large rock outcrop on which I observed two *A. aeneus*. I found other smaller boulders with some crevices at this site but no other large outcroppings. I confirmed occupation in subsequent visits, designating it as a new locality for the species.

The positive site from Transylvania County was predicted present by model 3. All other models predicted this site as absent. MaxEnt model 3 was therefore the only model that predicted presence at my newly discovered *A. aeneus* site.

## DISCUSSION

My models indicated that forested areas, intermediate elevations, and shallow soils of particular types are desirable landscape features for *A. aeneus*. Soil type was the most important predictor variable in our models. These models were built based on green salamanders found in rock crevices and it is no surprise that soil attributes are important predictors of rock outcrop habitat. Still, the single parameter of depth-to-bedrock was not as robust a predictor as a soil category, suggesting that other attributes of soil also determine quality habitat for *Aneides aeneus*. Canopy cover and elevation also played roles in predicting distribution indicating that the microclimate surrounding rocks is important in predicting green salamander habitat. My models predicted the highest likelihood of presence along the drop of the Blue Ridge Escarpment, and in areas within this region where soil and forest cover are optimal in the event an outcrop is available.

MaxEnt models performed very well in predicting suitable habitat for *A. aeneus*, however, landscape attributes are only one component in habitat suitability for a plethodontid salamander. Microenvironments within the landscape are also important, and depend on both landscape-level and fine-resolution factors. Presence of rock outcrops with specific crevice features may be a limiting factor for establishment of a green salamander population. My models lacked the ability to definitively determine the presence of a rock outcrop. Despite the lack of microhabitat detail, landscape models were useful in identifying potentially important habitat conditions within which there may exist rock outcroppings suitable for *A. aeneus* populations.

These models helped discover a new unknown location for *A. aeneus*. Two more previously unknown locations were identified when a land trust used these models to survey for *A. aeneus* (Kyle Pursel pers comm). Most areas we search did not have populations present but there are several candidate areas on private land left to be surveyed.

Another important component to distributions of species, especially of amphibians, is a locations' land use history. Western North Carolina has an extensive history of anthropogenic disturbance including heavy logging and large fires. Some of my locations seemed to have good habitat, however, past land use may have caused local extirpations and re-colonization may have been slow or impeded.

There are many methods to choose from when given the task of modeling the distribution of a species (Guisan and Thuiller 2005; Araujo and Guisan 2006; Hirzel and Le Lay 2008; Newbold 2010). My models represent easy access algorithms for a conservation planner to use when determining priority areas for a given species. Analyses were based on a presence-only dataset, a common scenario for conservation biologists. These results showed that MaxEnt, using NRCS soils data, was a useful way to quickly find new green salamander locations.

An advantage of MaxEnt was the ability to incorporate a categorical variable with several hundred categories. Keeping the soils layer in its original NRCS categories made a tremendous difference in model outcome. Depth-to-bedrock appeared to be important in the model and it made sense it would play a role in the presence of rock-outcroppings. My first MaxEnt model used this continuous variable and while outperforming the logistic regression, ranked behind both MaxEnt models using soil-type categories.

This suggests that soil depth is not the only soil attribute important for *A. aeneus* habitat. One avenue to explore would be to model several more attributes in order to discover which of these together might better predict *A. aeneus*.

When applied to other parts of *A. aeneus* range, depth-to-bedrock may prove to be more important. In the Cumberland Plateau, extensive rock walls are more common and home to large populations of green salamanders. The Blue Ridge escarpment does not contain these same geologic structures but rather has smaller more dispersed outcroppings that may be less connected to main bedrock layers. Bedrock exposure may not be as predictive of presence as maybe in the Cumberland populations.

Addition of error rates in model evaluations improved comparisons. Both commission and omission rates were important in determining the usefulness of each model and revealed a marked difference between MaxEnt and logistic regression models. Commission rates can be as or even more important than omission rates (Rondinini et al. 2006), and addition of these helped define specificity of the models. My study had the goal to decrease search area, therefore, low commission is particularly preferred. One study used both true absences and pseudo-absences to evaluate error rates on models built on pseudo-absences. They found error rates calculated on pseudo-absences correlated with, but underestimated, true commission error rates (Vaclavik and Meentemeyer 2009). Logistic regressions may, therefore, have even higher commission rates and are less specific in determining what makes a green salamander absent.

Model 3 was the only model to predict presence of *A. aeneus* at my newly discovered site. The other models predicted presence in many raster cells in the area around that site, but only Maxent model 3, using all covariates and soil as a category, was

correct in predicting a presence at that particular pixel. I urge modelers to incorporate measures such as error rates alongside AUC to evaluate their models, and other studies have come to similar conclusions (Lobo et al. 2007; Smith 2013)

The MaxEnt program created reliable models that can be used easily in a GIS interface. This algorithm, available in a very user friendly program, may be unjustly viewed as inferior because of the ease of access to potentially ill-advised modelers. However, it is not the fault of the algorithm if it is used incorrectly. Species distribution models are becoming an important component of wildlife conservation (Pearce and Ferrier 2000; Elith and Leathwick 2009), and easy accessibility of MaxEnt may have helped develop that trend. Furthermore, many agencies that need quick results for immediate decisions are dealing with presence-only datasets and do not have the option to create presence/absence models.

My models were useful in determining where to survey. There has already been great search effort for *A. aeneus* on government-owned lands especially in areas that fall within the known range. Private land, on the other hand, offers more area with less history of intensive surveying and where these models can be of most use. Conservation planners may suspect animals on these private lands and models of strong predicted presence may provide more incentive to consult private landowners and collect more detailed information on which properties to gain access to for future surveys and potential easements.

These models predicted the availability of potential habitat conditions at the landscape level habitat required for *Aneides aeneus* and I am confident they will be valuable in establishment of priority areas for the species. Presence-only methods using



machine learning algorithms like MaxEnt give modelers a reliable option when choosing from available datasets. While presence-only datasets do not supply detection rates or unbiased data, they can be a good alternative especially when dealing with species where detection is high and consistent. The green salamander provides an example of such a species and demonstrates usefulness of presence-only modeling for priority species.

MaxEnt models are a starting point to help narrow the search for biologists interested in establishing priority areas for *A. aeneus*.

## LITERATURE CITED

- Aguirre-Gutiérrez, J., Carvalheiro, L.G., Polce, C., van Loon, E.E., Raes, N., Reemer, M., Biesmeijer, J.C. 2013. Fit-for-purpose: species distribution model performance depends on evaluation criteria – dutch hoverflies as a case study. *PLoS ONE* 8 e63708.
- Araújo, M.B. and Guisan, A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography* 33:1677–88.
- Austin, M.P. 2002. Spatial prediction of species distribution an interface between ecological theory and statistical modeling. *Ecological Modelling* 157:101-118.
- Austin, M.P., Gaywood, M.J. 1994. Current problems of environmental gradients and species response curves in relation to continuum theory. *Journal of Vegetation Science* 5:473-482.
- Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* 3:327–338.
- Brimley, C.S. 1944. Amphibians and reptiles of North Carolina. Carolina Tips, Number 7, Elon College, North Carolina.

- Bruce, R.C. 1968. The Role of the Blue Ridge Embayment on the Zoogeography of the Green Salamander, *Aneides aeneus*. *Herpetologica* 24:185-194.
- Corser, J.D. 2001. Decline of disjunct green salamander (*Aneides aeneus*) populations in the southern Appalachians. *Biological Conservation* 97:119-126.
- Elith, J. and Leathwick, J.R. 2007. Predicting species' distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Diversity and Distributions* 13:165-175.
- Elith, J. and Leathwick, J.R. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics* 40:677-697.
- Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. McC., Peterson, A.T., Phillips, S. J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R. E., Soberón, J., Williams, S., Wisz, M.S. and Zimmermann, N.E. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129-151.

- Engler, R., Guisan, A.H., Rechsteiner, L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology* 41:263-274.
- Eschtruth, A.K., Cleavitt, N.L., Battles, J.J., Evans, R.A., Fahey, T.J. 2006. Vegetation dynamics in declining eastern hemlock stands: 9 years of forest response to hemlock woolly adelgid infestation. *Canadian Journal of Forest Resources* 36:1435-1450.
- Franklin, J. 2009. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge, UK Cambridge University Press 338 p.
- Gordon, R.E. 1952. A contribution to the life history and ecology of the plethodontid salamander *Aneides aeneus* (Cope and Packard). *American Midland Naturalist*, 47:666-701.
- Gordon, R.E. 1967. *Aneides aeneus*. Pp. 30.1–30.2. Catalogue of American Amphibians and Reptiles. Society for the Study of Amphibians and Reptiles, St. Louis, Missouri.
- Guisan, A., Thuiller, W. 2005. Predicting species distribution: Offering more than simple habitat models. *Ecology Letters* 8:993-1009.

- Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P., Buckley, Y.M. 2013. Predicting species distribution for conservation decisions. *Ecology Letters* 16:1424-1435.
- Highton, R.T. 2005. In: Lannoo, M. (Ed.), *Amphibian Declines: The Conservation Status of United States Species*. University of California, Berkeley.
- Hirzel, A.H. and Le Lay, G. 2008. Habitat suitability modeling and niche theory. *Journal of Applied Ecology* 45:1372–81.
- Hutchinson, G.E. 1957. Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22:415–427.
- Jiménez-Valverde, A., Lobo, J.M. 2006. The ghost of unbalanced species distribution data in geographic model predictions. *Diversity and Distributions* 12:521–524.
- Lipps, G.J. 2005. A framework for predicting the occurrence of rare amphibians. A case study with the green salamander. *Thesis* Bowling Green State University 52 pp.

- Liu, C.R., Berry, P.M., Dawson T.P., Pearson R.G. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28:385–393.
- Lobo, M.J., Jimenez-Valverde, A., Hortal, J. 2010. The uncertain nature of absences and their importance in species distribution modeling. *Ecography* 33:103-114.
- Lobo, M.J., Jimenez-Valverde, A., Real, R. 2007. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* 17:145-151.
- Miloski, S.E. 2010. Movement patterns and artificial cover use of green salamanders (*Aneides aeneus*) in Kanawha county, West Virginia. *Thesis* Marshall University
- Netting, M.G., Richmond, N. 1932 The green salamander, *Aneides aeneus*, in northern West Virginia. *Copeia* 1932:101-102.
- Newbold, T. 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models *Progress in Physical Geography* 34:3-22.
- Pandit, S.N., Hayward. A., Leeuw, J. and Kolasa. J. 2010. Does plot size affect the performance of GIS-based species distribution models? *Journal of Geographical System* 12:389-407.

- Pauley, T.K., Kees, G., and Pollio, C. 1993. Status of the green salamander (*Aneides aeneus*) in the New River Gorge National River. *Proceedings of the West Virginia Academy of Science* 65:48.
- Pearce, J., Ferrier, S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133:225-245.
- Petranka, J.W. 1998. *Salamanders of the United States and Canada*. Smithsonian Institution Press, Washington, DC.
- Petranka, J.W., Eldridge, M. E., Haley, K. E. 1993. Effects of timber harvesting on southern Appalachian salamanders. *Conservation Biology* 7:363–377.
- Peterson, 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography* 30:550-560.
- Phillips, S.J., Anderson, R.P., Schapire, R.E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190:231-259.
- Raes, N., ter Steege, H. 2007. A null model for significance testing of presence-only species distribution models. *Ecography* 30:727-736.

- Rebelo, H., Jones, R. 2010. Ground validation of presence-only modelling with rare species: a case study on barbastellus *Barbastella barbastellus* (Chiroptera: Vespertilionidae). *Journal of Applied Ecology* 47:410-420.
- Reed, R.C Jr., Hicks, J., Williams, L.A., Patch, S.C. 2009. Attributes of rock crevices selected by the green salamander, *Aneides aeneus*, on the Blue Ridge Escarpment. *Herpetological Review* 40:151-153.
- Renner, I.W., Warton, D.I. 2013. Equivalence of MAXENT and poisson point process models for species distribution modeling in ecology. *Biometrics* 69 274-281
- Riedel, B., Russell, K.R., Ford, W., Godwin, H.W. 2006. *Aneides aeneus* (Green Salamander). Dispersal. *Herpetological Review* 27:196-197.
- Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. 2006. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters* 9:1136–1145.
- Rovito, S.M., Parra-Olea, G., Vasquez-Almazán, C.R., Papenfuss, T.J., Wake, D.B. 2009. Dramatic declines in neotropical salamander populations are an important part of the global amphibian crisis. *Proc. Natl. Acad. Sci.* 106:3231–3236.



- Smith, A.B. 2013. On evaluating species distribution models with random background sites in place of absences when test presences disproportionately sample suitable habitat. *Diversity and Distributions* 19:867-872
- Snyder, D.H. 1983. The apparent crash and possible extinction of the green salamander, *Aneides aeneus*, in the Carolinas. *Association of the Southeastern Biologists Bulletin* 30 (abstract)
- Stuart, S.N., Chanson, J.S., Cox, N.A., Young, B.E., Rodrigues, A.S.L., Fischman, D.L., Waller, R.W. 2004. Status and trends of amphibian declines and extinctions worldwide. *Science* 306:1783-1786.
- Syfert, M.M., Smith, M.J., Coomes, D.A. 2013. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PLoS ONE* 8:e55158.
- Vaclavik, T., Meentemeyer, R. K. 2009. Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling* 22:3248-3258.
- Wake, D.B., Vredenburg, V.T. 2008. Colloquium paper: Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proc Natl Acad Sci* 105:11466–11473.

- Waldron, J., Humphries, J. 2005. Arboreal Habitat Use by the Green Salamander, *Aneides aeneus*, in South Carolina. *Journal of Herpetology* 39:486-492
- Warren, D.L., Seifert, S.N. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications* 21:335-342.
- Wilson, C.R. 2003. Woody and arboreal habitats of the Green salamander *Aneides aeneus* in the Blue Ridge mountains. *Contemporary Herpetology* 2
- Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, A., Nichols, J.D., Grant, E.C., Veran, S. 2013. Presence-only modeling using MAXENT: when can we trust the inferences? *Methods in Ecology and Evolution* 4:236-243.
- Zaniewski, A.E., Lehman, A., Overton, J.M. 2002. Predicting species spatial distributions using presence-only data: A case study of native New Zealand ferns. *Ecological Modelling* 157:261-280.